Aalto University
School of Business

# Corporate sustainability and stock returns: Evidence from a global dataset

Jukka Song

Bachelor's Thesis, Finance

December 5, 2016

## Abstract

This paper analyzes the relationship between corporate sustainability and long-run stock returns. A value-weighted portfolio of the "Global 100 Most Sustainable Corporations in the World" lost an annual four-factor alpha of -4.3% from 2005 to 2016. The results are robust to controlling for different weighting methodologies, asset pricing models, comparing to industry-adjusted benchmarks, and geographical differences. These findings suggest overvaluation of environmental, social, and governance (ESG) factors by the market, exhibited in the globally highest percentile of sustainable firms in their respective industries. For investors, these findings suggest using aggregate ESG ratings-based socially responsible investing (SRI) screens may dent their investment returns.

Department of Finance
Aalto University School of Business
Thesis supervisor: Matthijs Lof

**Author** Jukka Song

**Title of thesis** Corporate sustainability and stock returns: Evidence from a global dataset

**Degree** Bachelor's degree

**Degree programme** Finance

**Thesis advisor(s)** Matthijs Lof

**Year of approval** 2016 **Number of pages** 25 **Language** English

**Abstract**

This paper analyzes the relationship between corporate sustainability and long-run stock returns. A value-weighted portfolio of the "Global 100 Most Sustainable Corporations in the World" lost an annual four-factor alpha of -4.3% from 2005 to 2016. The results are robust to controlling for different weighting methodologies, asset pricing models, comparing to industry-adjusted benchmarks, and geographical differences. These findings suggest overvaluation of environmental, social, and governance (ESG) factors by the market, exhibited in the globally highest percentile of sustainable firms in their respective industries. For investors, these findings suggest using aggregate ESG ratings-based socially responsible investing (SRI) screens may dent their investment returns.

**Keywords** sustainable, investing, socially, responsible, corporate, sustainability

## Contents

## 1. Introduction

This paper analyzes the relationship between corporate sustainability and long-run stock returns. A value-weighted portfolio of the "Global 100 Most Sustainable Corporations in the World"[1] lost a four-factor alpha of -0.36% per month from 2005 to 2016, or -4.3% per year. The results are robust to controlling for different weighting methodologies, asset pricing models, comparing to industry-adjusted benchmarks, and geographical differences. For example, the equal-weighted portfolio underperformed by an alpha of -0.21% per month, or -2.5% per year. When compared to industry-matched benchmarks, the alpha worsens to a statistically significant -5.9% per year. When examining for geographical differences in the two main regions, US and Europe, I find that for a portfolio limited to only US firms, the alpha was -5.2% per year, and for European firms the alpha was insignificant. These findings contribute to current research concerning socially responsible investing (SRI). Specifically, the effect of SRI screens on investment performance.

Socially responsible investing is an investment discipline that considers environmental, social, and governance (ESG) criteria, in addition to traditional financial metrics, in investment decisions. The discipline has many terms and labels coined to describe it, such as sustainable, responsible, impact, green, ethical, values-based, or community investing, but they all refer directly or partially to some aspect of the ever more maturing and widespread investment discipline that includes ESG factors in conducting investment decisions. There are various approaches to SRI, with a common one being negative or positive screening. In negative screening, an investor excludes individual companies, industries or countries from portfolios if their areas of activity conflict with an investor's value. The opposite variant of this is positive screening (also called best-in-class screening), where investments are focused into companies or industries with higher ESG performance relative to peers.[2] SRI has attracted considerable research interest, due to the significant growth the industry has seen: the total amount of SRI assets in the United States has grown from around $700 billion in 1995 to $6.57 trillion in 2014, representing a total growth of 929% over the years.[3]

---

[1] Found from the Corporate Knights's website: http://www.corporateknights.com/reports/global-100/

[2] Example of SRI practices from global leading bank, UBS: https://www.ubs.com/content/dam/WealthManagementAmericas/documents/adding-values-to-investing-2015-03-sustainable-investing.pdf

[3] US SIF Report on SRI trends: http://www.ussif.org/Files/Publications/SIF_Trends_14.F.ES.pdf

### *1.1. Literature review and theoretical motivations*

Although conventional investment firms are "increasingly active in creating and marketing targeted products for sustainable investors"[4], finance research has yet to reach a conclusion on the financial profitability of the discipline. There are considerable findings supporting both sides of the argument, both theoretical and empirical. For example, negative screens are argued to theoretically decrease returns, because of lower diversification (Markowitz, 1959) and the shunned-stock hypothesis (Hong and Kacperczyk, 2009; Statman and Glushkov, 2009; Derwall, Koedijk, and Ter Horst, 2011), while positive screens have found to be able to yield superior returns due to undervalued intangibles (Edmans, 2011; Gompers, Ishii, and Metrick, 2003; Deng, Kang and Low, 2013; Derwall et. al., 2011; Derwall, Guenster, Bauer and Koedijk, 2005; Statman and Glushkov, 2009; Cremers and Nair, 2005; Galema, Plantinga and Scholtens, 2008; Kempf and Osthoff, 2007). For investors, the key question is whether ESG factors are valued correctly by the market. If ESG factors significantly improve firm performance but are undervalued, SRI should result in positive risk-adjusted excess returns. If, however, ESG factors don't significantly improve firm performance but are overvalued, SRI should result in negative excess returns.

Research findings documenting negative SRI returns center around three reasons for the results: shunned-stock hypothesis, limited optimization argument and overvaluation. The shunned-stock hypothesis posits that socially controversial stocks have superior returns because they are shunned by ethical investors who push their prices below those of responsible stocks, all else equal. (Derwall et. al., 2011; Heinkel, Kraus, and Zechner, 2001) Thus, socially responsible investors, especially institutions who are subject to norms, "pay a financial cost in abstaining from these stocks" (e.g. tobacco, weapons, alcohol companies). (Hong and Kacperczyk, 2009; Statman and Glushkov, 2009) For limited optimization, the Markowitz (1959) argument suggests that any SRI screen worsens performance, since according to traditional portfolio theory restricting an investor's choice set results in a constrained optimization that is never better than an unconstrained optimization, mathematically. This argument holds particularly strong, with ESG factors that have no clear consensus on their financial desirability. In such a case, investors are made unambiguously worse off by the screening-out process, as removing stocks on ethical grounds from the investable universe will reduce portfolio efficiency. (Brammer, Brooks, and Pavelin, 2006)

[4] US SIF Report of SRI trends: http://www.ussif.org/Files/Publications/SIF_Trends_14.F.ES.pdf

For overvaluation, various hypotheses have been made on to why ESG factors might be overvalued. First, behavioral finance arguments center on values-based[5] investors, who consciously accept paying a price for investing according to their ethical values (Renneboog, Ter Horst, and Zhang, 2008; Derwall et. al., 2011). For example, an institutional investor vulnerable to societal norms and public opinion may be willing to forgo financial returns to instead receive a non-financial payoff of upholding and encouraging socially responsible practices. (Hong and Kacperczyk, 2009) Second, the errors-in-expectations hypothesis posits that the future tangible benefits that certain ESG factors might generate are overestimated by the market. (Statman and Glushkov, 2009). For example, due to mounting research findings in support of the positive financial benefits of corporate social responsibility (CSR) practices (Margolish, Elfenbein, and Walsh, 2007; Orlitzky, Schmidt, and Rynes, 2003; Flammer, 2015) the market expects such practices to improve firm performance. However, due to the difficult nature of such intangible valuation, if the real benefits fall short of expectations, the stocks are likely to underperform. Third, theoretically if a significantly large proportion of the investor base adheres to SRI principles, list inclusion itself could drive the prices of SRI stocks above their fundamental values. For example, list inclusion and improvement of ESG performance of one company may allow the company to pass certain SRI screens of institutional investors. A large movement of such demand would then drive up the stocks prices. However, this argument has so far not found empirical proof.

Research findings documenting positive returns center around results showing undervaluation of certain ESG factors. Undervaluation of ESG factors may happen, because the costs of social responsibility are immediate, but benefits are less obvious and in the future (Statman and Glushkov, 2009). Research findings that show ESG factors yielding excess returns suggest that there are individual ESG factors that improve firm performance, but are not immediately valued by investors, until the benefits manifest in subsequent tangible outcomes. Also, there is ample previous evidence that the market fails to fully incorporate intangibles in general, such as employee satisfaction (Edmans, 2011), R&D (Lev and Sougiannis, 1996; Chan, Lakonishok, and Sougiannis, 2001), advertising (Chan, Lakonishok, and Sougiannis, 2001), patent citations (Deng, Lev, and Narin, 1999), and software development costs (Aboody and Lev, 1998), that all earn superior long-run returns. Then, the market may be even more likely

[5] The term refers to any social investment that is rooted in non-pecuniary motivations, but instead social or personal values.

to undervalue ESG factors since theory has ambiguous predictions for whether they are desirable for firm value.

Examples of ESG factors that have been shown to improve stock returns are employee satisfaction (Edmans, 2011), KLD[6] corporate social responsibility scores (Deng et. al., 2013), shareholder rights (Gompers et. al., 2003), Innovest's intangible value assessment ratings (Aktas, de Bodt, and Cousin, 2011), and KLD subcategory community support (Statman and Glushkov, 2009). For specific time periods, long-short portfolios based on aggregate KLD sustainability ratings have also been found to yield abnormal positive returns. (Kempf and Osthoff, 2007) Other aspects of SRI not directly related to superior stock returns are better performance of SRI funds during market crises (Nofsinger and Varma, 2014) and SRI fund investors exhibiting higher resilience than investors in conventional mutual funds, suggesting SRI investors move money in and out of their mutual funds at slower rates than investors in other funds. (Bollen, 2007; Benson and Humphrey, 2008; Bialkowski and Starks, 2016) Additionally, the limited optimization argument has been challenged by Edmans (2011), who states: "If an investor is aware of every asset in the economy, an SRI screen may never help, as non-SRI investors are free to choose the screen stocks anyway. However, if she can only learn about a subset of the available universe due to time constraints (Merton, 1987), the SRI screen – rather than excluding good investments – may focus the choice set on good investments." This argument lies on the assumption that ESG factors that SRI screens consider are value-relevant.

Research findings have also shown mixed or no-effect results for SRI. Aggregate measures of SRI have been shown to possibly confound existing relationships between individual dimensions of SRI and returns. (Galema et. al., 2008) Derwall et. al. (2011) argue that values-driven SRI investors using unprofitable negative screens and profit-seeking SRI investors who use profitable positive screens co-exist, and that for most SRI strategies incorporating both negative and positive screens, the gains of positive screening are counteracted by the losses of negative screening. Statman and Glushkov's (2009) findings support this argument, where on average, aggregate ESG ratings consist of both profitable and unprofitable ESG factors, and thus SRI funds using aggregate ESG measures often return zero or negative excess returns. (Renneboog et. al., 2008) Additionally, benefits of some ESG factors that yielded abnormal returns in one period have later been shown to have lost such returns. (Edmans, 2011; Core,

[6] KLD stands for Kinder, Lydenberg and Domini. The company KLD Research & Analytics Inc. was a leading provider of ESG research for institutional investors. The company was acquired by MSCI in 2010, and is now known as MSCI ESG Research.

Guay, and Rusticus, 2006; Bebchuk, Cohen, and Wang, 2013) This further decreases the likelihood of persistent abnormal returns gained through SRI. So, due to theory and empirical evidence having support for both positive and negative excess returns, no-effect results may either indicate that markets are correctly valuing ESG factors, or that opposite effects counteract each other depending on the research method.

### *1.2. Hypothesis and key question*

Evidently, theory yields very conflicting predictions on the profitability of SRI, shown by the variance of different empirical results found in prior research. Since my data uses an aggregate measure of ESG performance, based on the discussion in the previous subsection, the hypothesis should be either no effect or overvaluation. To focus my research, my hypothesis is:

H1: Environmental, social, and governance (ESG) factors are overvalued, so the alpha from an ESG-performance-sorted stock portfolio is negative.

The key question in this paper is whether the market values ESG factors correctly, as this is one of the most important problems for investors considering SRI. To investigate this question and test for the hypotheses, using the stock returns of the Corporate Knight's "Global 100" list of global top 100 sustainable companies should be insightful. This is because the Global 100 constituents should, by definition, be the ones with the strongest ESG performance globally, in their respective industries. If ESG factors improve firm performance, any such benefits should then be the most pronounced in the Global 100. The benefits should also be among the most well-communicated, so the nature of the data also mitigates saliency concerns. Thus, the stock returns of a portfolio formed from these firms can be used to determine whether ESG factors are over- or undervalued. If the Global 100 portfolio produces negative excess returns, it would suggest that ESG factors are overvalued. In other words, the costs of sustainability exceed future realizable benefits, but the market is over-optimistic about these benefits.

## 2. Data and summary statistics

My main data source is the list of "Global 100 Most Sustainable Corporations in the World"[7], hereon referred to as the Global 100 list. The list has been published annually since 2005 by Corporate Knights, a Toronto-based, specialized media and investment research firm. The Global 100 list is published each year in late January at the World Economic Forum and in leading media including Forbes.com, as well as the Corporate Knights own website. I form an annually rebalanced portfolio of stocks based on the list of companies in the Global 100 list, starting from February 1, in order to track the stock returns of the globally most sustainable companies for the year following their list inclusion.

The methodology of formulating the Global 100 list is quantitative and based on publicly-disclosed data. The methodology consists of a series of initial negative screens to form a shortlist of firms, which are then ranked based on up to twelve key performance indicators (KPI). First, the starting universe is lined out to be all companies that had a market capitalization of $US 2 billion. Market capitalization data is taken each year on October 1st. Second, the starting universe companies are tested through four screens: sustainability disclosure, financial health, sin screen, and sanctions. In the sustainability disclosure screen, companies that did not disclose at least 75% of their "priority KPIs" in their respective GICS industry group are eliminated. A "priority KPI" here refers to any of the 12 KPIs that is disclosed by at least 10% of all companies in a given industry group. In the financial health screen, a Piotroski F-Score test[8] is conducted for each firm, where firms scoring 4 or less are eliminated. In the sin screen, companies with a GICS Sub-Industry classification equal to "Tobacco" are eliminated. Companies with classification of "Aerospace & Defense" are revenue tested: if a company derives a majority of its revenue from its Defense business group (e.g. weapons manufacturing), it is eliminated. In the sanctions screen, companies are examined based on their dollar amount paid on a trailing one-year basis in sustainability-related fines, penalties or settlements. If the total amount of such sanctions for a company as a percentage of total revenue during the 12 months is found to be in the bottom quartile compared to GICS Industry Group peers, the

[7] Found on the Corporate Knights's website: http://www.corporateknights.com/reports/global-100/

[8] The Piotroski F-Score test consists of nine individual tests, where each test scores one for pass and zero for fail. The tests are: 1) net profit is positive 2) operating cash flow is positive 3) net profit/total assets at beginning of year, minus the same number for the previous year is positive 4) operating cash flow is greater than net profit 5) long term debt/average assets has not increased 6) the current ratio has increased 7) no raising of ordinary (common) equity over the previous year 8) gross margin has improved over the previous year 9) asset turnover has increased.

company is eliminated. In addition to the four screens, any company that was on the previous year's Global 100 list is automatically included in the current year's shortlist if it is not in the bottom quartile of the sanctions screen.

Third, after forming the shortlist, the 12 KPIs for which the Global 100 is known are introduced, shown in Table 1. All companies on the shortlist are scored on a percent rank basis against their global industry peers on the "priority KPIs" (any of the 12 KPIs that is disclosed by at least 10% of all companies in a given industry group) for their respective GISC Industry Group. Each company is then assigned an overall score that is an average of the scores of each priority KPI. Finally, the Global 100 list industry composition is matched against its benchmark MSCI ACWI, by assigning a fixed number of slots for each GICS Sector equaling the sector's value-weight in MSCI ACWI (e.g. if 10% of the MSCI ACWI consists of Financial sector companies, 10 slots in the Global 100 would be reserved for Financial companies). Subject to these slot assignments, the Global 100 list is then filled with companies receiving the top overall score in their respective GISC Sectors. This way, the list represents companies that have the globally highest relative ESG performance compared to their global industry peers.

For this paper, I form the annually rebalanced Global 100 portfolio starting from February 1, 2005 (the first Global 100 list was published January 22, 2005). The companies included in the portfolio are exactly the companies that are on the Global 100 list each year. I retrieve the stock prices of each firm for the first day of each month, from February 1, year $t$, to February 1, year $t + 1$. This way, I can calculate the monthly stock returns for a year, until a new Global 100 is published again the following year. The portfolio is then reformed based on the following year's Global 100 list, and this process is repeated until September 1, 2016 (the last available monthly stock data date). The stock price data is retrieved from Datastream, where I individually searched each firm by name, and matched it with its Datastream code to search for stock prices. I chose B-class stocks when there were more than one equity class to choose from. When choosing a stock with multiple stock exchange listings, I chose the stock exchange the company was headquartered in. While retrieving the stock price data, I extracted the price data in US dollars directly using Datastream's currency conversion function that automatically returns the correct prices in the desired currency. This way, the returns are fully comparable despite local currency differences.

**Table 1.** Twelve key performance indicators (KPI) for ranking corporate sustainability in the Global 100 list formation process.[9]

| KPI | Methodology |
|---|---|
| *Resource management* | |
| Energy productivity | *Revenue / energy use* |
| Carbon productivity | *Revenue / GHG emissions* |
| Water productivity | *Revenue / water withdrawal* |
| Waste productivity | *Revenue / non-recycled or reused waste generated* |
| *Financial management* | |
| Innovation capacity | *R&D expenses / revenue* |
| Percentage tax paid | *Cash tax / EBITDA* |
| CEO-average employee pay | *CEO compensation / average employee compensation* |
| Pension fund status | *Unfunded liabilities / total assets* |
| *Employee management* | |
| Safety performance | *Fatalities and lost time incidents* |
| Employee turnover | *Number of departures / average total employees* |
| Leadership diversity | *Female representation on board of directors and executive management team* |
| Clean capitalism pay link | *Mechanisms that link senior executive pay to clean capitalism targets* |

I form both equal- and value-weighted portfolios to account for possible differences of results depending on weighting methodologies (Fama and French, 2008). For the equal-weighted portfolio, I calculate the average monthly return for the portfolio as the average of monthly returns for all the stocks in the portfolio at the time. For the value-weighted portfolio, I calculate each company's stock value-weight as the company's market capitalization (taken from Datastream) divided by the sum of market capitalizations of the companies in the portfolio. I multiply the monthly returns of the companies with their respective value-weights, and sum the monthly value-weighted returns of all companies to receive the value-weighted monthly return of the portfolio.

Table 2 details the number of companies each year that had stock returns available on Datastream. There were mostly 100 companies each year, with some missing companies being due to empty lines on the original Global 100 list, or unavailable data on Datastream. The table also shows the number of firms added and dropped each year: on average 35 added and dropped each year. Intuitively, corporate sustainability can be a relatively persistent characteristic: a firm

[9] For a more detailed breakdown, see source website. Retrieved from the Corporate Knights's website: http://www.corporateknights.com/wp-content/uploads/2015/01/2015Global100_Methodology.pdf

that adopts sustainable practices most likely want to hold on to these practices consistently. However, it isn't permanent and is subject to external and internal changes, which shows in approximately a third of the firms dropping out each year.

**Table 2.** Summary statistics.

The second column details the number of the Global 100 Most Sustainable Corporations in the World that were included in each year of the Global 100 portfolio. The year's with less than 100 companies had missing data in the original Global 100 list, such as empty lines (2013), or stock price data could not be found from Datastream (2005 and 2015). The third column gives the number of new unique firms added to the portfolio of that year. The fourth column shows the number of firms on the previous year's portfolio which no longer feature in the current one.

| Year of portfolio | Number of companies | Added | Dropped |
|---|---|---|---|
| 2005 | 99 | | |
| 2006 | 100 | 28 | 27 |
| 2007 | 100 | 35 | 35 |
| 2008 | 100 | 32 | 32 |
| 2009 | 100 | 36 | 36 |
| 2010 | 100 | 51 | 51 |
| 2011 | 100 | 34 | 34 |
| 2012 | 100 | 32 | 32 |
| 2013 | 96 | 45 | 48 |
| 2014 | 100 | 33 | 29 |
| 2015 | 99 | 35 | 36 |
| 2016 | 100 | 26 | 25 |

To gain some insight on the geographical distribution of the Global 100 list, table 3 shows the average geographical distribution of the headquarters of companies. As can be seen from the table, the Global 100 list is heavily skewed to Europe and the United States, with Europe representing 56 of the 100 companies on average. The next heaviest representation goes to the United States, with an average of 18 out of 100 companies of the Global 100 headquartering there. The third spot is held by Japan, with an average of 9 out of 100 companies headquartered there. The Kenneth French "Global 3 Factors for Developed Markets" dataset that is used in analysis of the stock returns, uses value-weighted regional market returns to compile the global market return. Compared to that, the geographical distribution of the Global 100 list is not identical, but fairly representative. The Kenneth French global dataset for market returns contains all the countries the Global 100 firms are headquartered in, and is also skewed towards Europe and the United States.

**Table 3.** Geographical distribution of the companies on the Global 100 lists.

The second column details the average number of companies headquartered in the country in question on the Global 100 lists. The third column shows the median. For each year, the country each firm headquarters in is documented, and the geographical distribution for each year can be counted. This table shows the total average count and median of headquarter country distribution of all 11 Global 100 lists from 2005 to 2016.

| Country | Average | Median |
|---|---|---|
| **United Kingdom** | 18.0 | 17.5 |
| **United States** | 16.1 | 17.5 |
| **Japan** | 8.8 | 7.5 |
| **Canada** | 7.5 | 7.0 |
| **France** | 7.0 | 7.0 |
| **Germany** | 5.8 | 5.5 |
| **Australia** | 4.8 | 5.0 |
| **Sweden** | 4.5 | 4.5 |
| **Finland** | 3.7 | 3.0 |
| **Switzerland** | 3.6 | 3.0 |
| **Denmark** | 2.8 | 3.0 |
| **Netherlands** | 2.8 | 3.0 |
| **Spain** | 2.7 | 2.5 |
| **Norway** | 2.1 | 2.0 |
| **Singapore** | 1.7 | 1.5 |
| **Brazil** | 1.4 | 1.5 |
| **South Korea** | 1.4 | 1.0 |
| **Belgium** | 1.1 | 1.0 |
| **Italy** | 0.8 | 1.0 |
| **Austria** | 0.5 | 0.0 |
| **Hong Kong (SAR), China** | 0.5 | 0.5 |
| **India** | 0.5 | 0.0 |
| **Portugal** | 0.5 | 0.0 |
| **Ireland** | 0.4 | 0.0 |
| **South Africa** | 0.3 | 0.0 |
| **China** | 0.2 | 0.0 |
| **Taiwan** | 0.1 | 0.0 |
| **Total** | 99.5 | 94.5 |

Table 4 documents the market capitalization of the Global 100 companies. Notably, the mean market capitalization for the companies is very large, around US$ 40 billion. As a comparison, the 80$^{th}$ percentile breakpoint for the Kenneth French US Research Breakpoints Data from NYSE in January 2010 is approximately US$ 9 billion. However, the median is

significantly smaller at US$ 20 billion, which signals the presence of some mega-cap stocks that skew the average higher than the median. This finding implies the methodology of Global 100 favors highly transparent large-cap firms that have resources to fully comply with the disclosure requirements of the methodology.

**Table 4**. Market capitalization averages.

The market capitalization summary statistics of the Global 100 portfolio companies. The data is retrieved from Datastream. The market values of companies are retrieved in US$ millions, and presented in US$ billions. The market value of all companies in all of the 11 Global 100 lists from 2005 to 2016 are summed and averaged. The number of observations includes duplicates, where one firm is present in multiple Global 100 lists over the years.

| | # obs | Mean | Median | Std. dev. | Min | Max |
|---|---|---|---|---|---|---|
| **Market Value of Company (US$ billions)** | 1194 | 40.3 | 20.2 | 53016.0 | 0.1 | 534.7 |

Table 5 displays the ten industries with the largest average proportion in the Global 100 constituent companies. Due to the methodology of forming the Global 100 list, where a specific number of slots were assigned for each industry group based on their proportions in the MSCI ACWI index, it was expected that certain industries would have more slots. The industry distribution is not heavily skewed toward any single industry, so the final results are unlikely to be significantly impacted by the industry distribution. The largest five industries are Banks (9%), Oil (6%), Utilities (6%), Drugs (6%), and Rtail (5%).[10]

[10] The industry definitions and their abbreviations are according to the Kenneth French 49-Industry Portfolio found on their website:
http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/det_49_ind_port.html

**Table 5.** Average industry distribution of Global 100 constituents.

The second column displays the average number of companies belonging to the industry in question each year. The average count is calculated as the average number of companies in a given industry, not their market capitalization. Industry definitions and their abbreviations are according to the Kenneth French 49-Industry Portfolio (see footnote 9). The sample period is 2005-2016.

| **Industry** | **Average Count** |
|---|---|
| Banks | 8.8 |
| Oil | 5.8 |
| Util | 5.8 |
| Drugs | 5.5 |
| Rtail | 4.9 |
| Telcm | 4.9 |
| Insur | 4.8 |
| Hshld | 4.8 |
| Chips | 4.2 |
| Autos | 4.0 |
| Other | 45.8 |
| **Total** | **99.3** |
| # of industries | 44 |

## 3. Analysis and results

To ensure that any outperformance of the Global 100 does not result from risk, I control for the four Carhart (1997) factors using

$$R_{it} = \alpha + \beta_{MKT} MKT_t + \beta_{HML} HML_t + \beta_{SMB} SMB_t + \beta_{MOM} MOM_t + \varepsilon_{it} \qquad (1)$$

where $R_{it}$ is the return on the Global 100 portfolio in month *t* in excess of the risk-free rate, taken as the US one month T-bill rate, from the Kenneth French website. $\alpha$ is an intercept that captures the abnormal risk-adjusted return. $MKT_t$, $HML_t$, $SMB_t$, and $MOM_t$ are the returns on the market, value, size, and momentum factors, taken from the Kenneth French website using global data. Standard errors are calculated using Newey and West (1987), which allows for $\varepsilon_{it}$ to be heteroskedastic and serially correlated.

### *3.1. Core results*

Table 6 presents the core results of the paper, for the entire 2005-2016 period. In coherence with my hypothesis, the Global 100 portfolio does not generate positive excess risk-adjusted returns, but a negative one instead over both weighting schemes. For equal-weighted returns, the alpha is -0.21% monthly (-2.5% annually) below the risk-free rate, significant at the 10% level (*p*-value is 0.078). For value-weighted returns, the alpha is -0.36% monthly (-4.3% annually), and significant at the 1% level (*p*-value is <0.001). The magnitude of alpha and thus overpricing is within the bounds of plausibility implied by previous studies that demonstrate abnormal returns. For example, SRI funds in the US, the UK, and in many continental European and Asia-Pacific countries have been found to underperform their Fama-French-Carhart benchmarks by -2.2% to -6.5%, according to the study by Renneboog et. al. (2008).

Based on the results, the two main significant explaining factors are market and size, both significant at the 1% level for the value-weighed portfolio; the coefficient for the market factor $\beta_{MKT}$ is 1.085, which indicates that the portfolio returns follow market returns very closely; the coefficient for the size factor, $\beta_{SMB}$ is -0.329, which indicates that the portfolio is biased towards large cap stocks, and a portion of the underperformance is explained by the lesser returns of large cap stock compared to small cap. The value and momentum factors are insignificant, so they fail to explain the portfolio returns. These results are reasonable, as we noted in the data earlier, how the Global 100 consisted of very large firms that tend to follow the market more, resulting in a close to 1.0 $\beta_{MKT}$. The Global 100 list formation methodology clearly is biased towards large capitalization companies.

### *3.2. Robustness tests*

I perform robustness tests to test for other asset pricing models, industry-adjusted returns, and any effect from geographical distribution. First, I test the robustness of the results by running the portfolio return regression compared to an industry-matched benchmark, using the 49-industry classification of Fama and French (1997). I do this by subtracting the monthly returns of an industry-matched portfolio from the monthly returns of the Global 100 portfolio, instead of subtracting the risk-free rate. To create the industry-matched portfolio, I first determine the industry weights of each year in the Global 100 portfolio. Then, I create a

**Table 6.** Risk-adjusted returns.

Monthly regressions of returns to a portfolio of the "Global 100 Most Sustainable Corporations in the World" on the four Carhart (1997) factors, *MKT*, *HML*, *SMB*, and *MOM*. The dependent variable is the portfolio return less the risk-free rate. Panel A contains equal-weighted returns and Panel B contains value-weighted returns. The alpha is the excess risk-adjusted return. *t*-Statistics are in parentheses. Standard errors are calculated using Newey and West (1987). The sample period is February 2005 – September 2016.

| Panel A: Equal-weighted | |
|---|---|
| $\alpha$ | -0.21* |
| | (-1.77) |
| $\beta_{MKT}$ | 1.09*** |
| | (27.82) |
| $\beta_{SMB}$ | -0.07 |
| | (-0.70) |
| $\beta_{HML}$ | 0.18 |
| | (1.64) |
| $\beta_{MOM}$ | -0.14*** |
| | (-3.61) |
| Panel B: Value-weighted | |
| $\alpha$ | -0.36*** |
| | (-3.42) |
| $\beta_{MKT}$ | 1.04*** |
| | (23.15) |
| $\beta_{SMB}$ | -0.33*** |
| | (-3.52) |
| $\beta_{HML}$ | 0.10 |
| | (0.81) |
| $\beta_{MOM}$ | -0.04 |
| | (-0.83) |
| $N$ | 140 |

*: Significant at the 10% level; **: Significant at the 5% level; ***: Significant at the 1% level.

portfolio with equal industry weights, with each industry return taken from Kenneth French's website 49-industry portfolio returns, where each industry's monthly return can be found. Although the industry returns are based on US data, due to the heavy weight of US firms in the Global 100 list, the industry returns can still arguably be used as the industry adjustment factor for robustness test purposes. This is done to ensure that outperformance is not due to any industry-specific periodical out- or underperformance. As shown in Table 7, the industry-adjustment resulted in even more negative excess returns of -0.51% monthly (-5.9% annually). Thus, the underperformance is not due to any industry-based periodical underperformance. Notably, the four Carhart (1997) factors are all insignificant in this regression, as the results are most likely impacted by the industry-adjustment applied to the portfolio returns.

**Table 7**. Risk-adjusted excess returns over the industry-benchmark.

Monthly returns of the Global 100 portfolio on the four Carhart (1997) factors, *MKT, HML, SMB,* and *MOM*. The dependent variable is the value-weight Global 100 portfolio return less the industry-matched portfolio return. *t*-Statistics are in parentheses. Standard errors are calculated using Newey and West (1987). The sample period is February 2005 – September 2016.

| Industry-adjusted | |
|---|---|
| $\alpha$ | -0.51*** |
| | (-3.13) |
| $\beta_{MKT}$ | 0.08 |
| | (1.50) |
| $\beta_{SMB}$ | -0.22 |
| | (-1.60) |
| $\beta_{HML}$ | 0.06 |
| | (0.46) |
| $\beta_{MOM}$ | 0.07 |
| | (1.26) |
| $N$ | 140 |

*: Significant at the 10% level; **: Significant at the 5% level; ***: Significant at the 1% level.

Second, I test for the robustness of results when subjected to the Fama and French (2016) five-factor regression, to test whether the profitability and investment factors could explain part of the underperformance of Global 100 stocks. I control for the five factors using

$$R_{it} = \alpha + \beta_{MKT}MKT_t + \beta_{HML}HML_t + \beta_{SMB}SMB_t + \beta_{RMW}RMW_t + \beta_{CMA}CMA_t + \varepsilon_{it} \quad (2)$$

where $R_{it}$ is the return on the Global 100 portfolio in month *t* in excess of the risk-free rate (US one month T-bill). $\alpha$ is an intercept that captures the abnormal risk-adjusted return. $MKT_t$, $HML_t$, $SMB_t$, $RMW_t$, $CMA_t$, are the returns on the market, value, size, profitability and investment factors, taken from the Kenneth French website using global data. Notably, the SMB factor from Kenneth French is calculated differently compared to the SMB factor used in the four-factor regression. Standard errors are calculated using Newey and West (1987), which allows for $\varepsilon_{it}$ to be heteroskedastic and serially correlated.

Table 8 documents the results of the Fama and French (2016) five-factor regression robustness test. For the value-weighted returns, the alpha is -0.36% monthly (-4.2% annually), and significant at the 1% level. When compared to the core results, the five-factor regression returned very similar results, with very similar alphas (-0.361% for three-factor model versus -0.358% for five-factor). The two main explaining factors are still market and size, both significant at the 1% level. The profitability and investment factors ended up insignificant in

the five-factor regression, which suggests that they failed to explain any of the underperformance. The core results are thus robust to different asset pricing models.

**Table 8.** Risk-adjusted excess returns using the Fama and French (2016) five-factor model.

Monthly regressions of returns to a portfolio of the Global 100 portfolio on the five Fama and French (2016) factors, *MKT, HML, SMB, RMW* and *CMA*. The dependent variable is the portfolio return less the risk-free rate. The Global 100 portfolio stock returns are value weighted. *t*-Statistics are in parentheses. Standard errors are calculated using Newey and West (1987). The sample period is February 2005 – September 2016.

| Fama-French 5-factor model | |
|---|---|
| $\alpha$ | -0.36*** |
| | (-3.30) |
| $\beta_{MKT}$ | 1.04*** |
| | (28.66) |
| $\beta_{SMB5f}$ | -0.34*** |
| | (-3.26) |
| $\beta_{HML}$ | 0.17 |
| | (1.15) |
| $\beta_{RMW}$ | -0.02 |
| | (-0.09) |
| $\beta_{CMA}$ | -0.03 |
| | (-0.13) |
| $N$ | 140 |

*: Significant at the 10% level; **: Significant at the 5% level; ***: Significant at the 1% level.

Third, I test for the robustness of results when isolated to only US or only Europe. The Global 100 portfolio is heavily skewed to US (average 18 out of 100 companies) and Europe (average 56 out of 100 companies), so examining the results when compared to region specific firms and benchmarks can be insightful to determine differences between the two main regions. To do this, I create US and Europe only value-weight portfolios and run the Carhart (1997) four-factor regressions, as in equation (1), to test them against their respective benchmarks taken from the Kenneth French website.[11]

Table 9 documents the regional portfolio risk-adjusted excess returns. For US, the alpha is -0.45% monthly (-5.2% annually), and significant at the 5% level. For Europe, the alpha is insignificantly different from zero. Compared to the core results, where annual alpha was -4.3%, the underperformance is stronger when considering only US firms. Comparing the core results to Europe, it seems that European Global 100 stock returns are less volatile than the

[11] The benchmark definitions can be found on the Kenneth French website at: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/f-f_3developed.html

market (0.91 beta). The results show a slight difference between the two regions in comparison to the global core results. In US; the results suggest and stronger mispricing, whereas in Europe ESG factors could be more correctly priced. Determining what drives the result in Europe would need further analysis. For example, it could be possible certain regions in Europe undervalue ESG factors and thus push the alpha closer to zero. Nonetheless, the alpha is still slightly negative, and shows some signs of overvaluation of ESG factors in European markets. Finding a definite answer to this is beyond the scope of this paper. Overall, the core results are robust to geographical adjustment.

**Table 9.** US and Europe specific risk-adjusted excess returns.

Monthly regressions of returns to region-specific Global 100 sub-portfolios on the four Carhart (1997) factors, *MKT, HML, SMB,* and *MOM.* The dependent variable is the portfolio return less the risk-free rate. Panel A contains value-weighted US returns and Panel B contains value-weighted European returns. The alpha is the excess risk-adjusted return. *t*-Statistics are in parentheses. Standard errors are calculated using Newey and West (1987). The sample period is February 2005 – September 2016.

| Panel A: US only | |
|---|---|
| $\alpha$ | -0.45** |
| | (-2.14) |
| $\beta_{USMKT}$ | 1.12*** |
| | (12.69) |
| $\beta_{USSMB}$ | -0.39*** |
| | (-4.85) |
| $\beta_{USHML}$ | 0.06 |
| | (0.48) |
| $\beta_{USMOM}$ | < 0.01 |
| | (0.05) |
| $N$ | 140 |
| Panel B: Europe only | |
| $\alpha$ | -0.08 |
| | (-0.71) |
| $\beta_{EUMKT}$ | 0.91*** |
| | (17.25) |
| $\beta_{EUSMB}$ | -0.43*** |
| | (-4.07) |
| $\beta_{EUHML}$ | -0.04 |
| | (-0.34) |
| $\beta_{EUMOM}$ | -0.04 |
| | (-0.95) |
| $N$ | 140 |

*: Significant at the 10% level; **: Significant at the 5% level; ***: Significant at the 1% level.

## 4. Discussion

The negative risk-adjusted excess returns found with the Global 100 portfolio supports the view that socially responsible investors pay a price for doing so, and that the market overvalues ESG factors. Because the Global 100 portfolio focuses on global stocks that should represent the most sustainable firms in relation to resource (energy, carbon, water, waste), financial (R&D, tax, CEO compensation, pension fund status), and employee (safety, turnover, diversity, executive pay) management, any financial benefits from such factors should be the most pronounced in these companies. The benefits should also be among the most well-communicated and widely known to investors, as suggested by the high disclosure requirements of the Global 100 list formation methodology. Then, finding the Global 100 portfolio to yield negative risk-adjusted excess returns would suggest that investors overvalue ESG factors. The analysis does not evaluate whether ESG factors actually improve firm performance, but only reveals how they are valued. Even if the ESG factors significantly improve firm performance, for example, as has been suggested by management literature on the positive relationship between corporate governance and firm valuation, (Beiner, Drobetz, Schmid, and Zimmermann, 2006; Eccles, Ioannou, and Serafeim, 2013) the negative alpha discovered in this paper suggest significant overvaluation of the real financial benefits received.

A portion of the negative alpha could stem from the negative effect related to limited optimization or shunned-stock hypothesis discussed in the introduction. Both of these effects are present in the Global 100 portfolio, as the methodology to form the Global 100 list excludes any company with market capitalization below US$ 2 billion from their starting universe, as well as excluding companies involved in the tobacco and weapons manufacturing business. Due to the exclusion of small cap firms, a large portion of potentially profitable stocks are excluded, and the mathematical optimization based on a limited choice set should be worse than an unlimited one. Due to the exclusion of tobacco and weapons manufacturing firms, potentially profitable stocks are ignored, as shown by Hong and Kacperczyk (2009). The analysis of this paper is limited in this regard, as I am unable to explore exactly how much of the underperformance stems from the two effects.

Another puzzle unsolved in my analysis is whether the overvaluation of Global 100 could stem from increased demand from SRI funds, due to list inclusion that now enables the firms to pass SRI screens. Edmans (2011) conducted a fund ownership analysis, and investigated whether SRI funds are overweighting the high employee satisfaction performers. If demand curves are downward-sloping, the increased buying by SRI funds could have led to increased stock prices. He found that institutions, banks, and insurance companies are all underweighted

on the “Best Companies” (high employee satisfaction performers). In addition, Heinkel et. al. (2001) found that exclusionary ethical investing leads to “polluting firms being held by fewer investors”, and that lack of risk sharing among non-green investors “leads to lower stock prices for polluting firms, thus raising their cost of capital”. Considering the opposite, when a significant proportion of a firm’s investors invest based on SRI principles, the opposite may occur, where cost of capital is lowered, and thus stock prices are higher. Based on these prior findings, it could be interesting for future research to investigate, whether institutional ownership on the Global 100 is overweight. Such a finding would support an explanation where list inclusion in the Global 100 increases the demand on the firms, and drives the prices over their fundamental values.

Due to lack of data, this paper did not conduct analysis on the individual effects of each of the twelve KPIs used in the methodology of forming the Global 100 list. It is likely that some of the KPIs yield a positive excess return, and some negative, as has been shown in similar research that break down aggregate ratings to their components (Statman and Glushkov, 2009; Derwall et. al., 2011). For future research, isolating the profitable ESG factors can be an important contribution specially to fund managers considering implementing SRI principles. Ideally, fund managers could selectively implement profitable SRI screens and achieve an SRI fund that yield superior returns while adhering to SRI principles.

Additionally, Edmans (2011) and Core, et. al. (2006) study earnings surprises of the companies performing high on some ESG factors and confirm that the superior returns of those companies stems partially from markets capitalizing on the value of ESG factors only after it manifests in tangible outcomes. For future research, it is possible that by studying the earnings surprises of the Global 100 stocks, one could find negative surprises. Such a finding would imply markets are over-optimistic about the benefits of ESG performance, and are later disappointed by the subpar tangible results. This would drive a portion of the observed underperformance of the Global 100 portfolio.

## 5. Conclusion

This paper finds that, globally, firms performing the highest on corporate sustainability in their respective industries generate negative risk-adjusted excess returns, even when controlling for industries, asset pricing models, or geographical differences. These findings imply that the market overvalues ESG factors, and lend weight to arguments suggesting investors pay a price in investing per SRI principles. The findings contradict, to an extent, earlier studies that show positive risk-adjusted excess returns for portfolios sorted on sustainability ratings (e.g. KLD ratings). (Statman and Glushkov, 2009; Kempf and Osthoff, 2007) The findings support the view that the use of aggregate ESG ratings, especially the mixed use of negative and positive screens that is common in SRI funds (Renneboog et. al., 2008), results in inferior returns. (Derwall et. al., 2011) While it is possible that individual ESG factors yield positive abnormal returns (Edmans, 2011), large aggregated indices or ratings most likely dilute or counteract any such benefits.

The unexplained underperformance found in this paper by the Carhart (1997) four-factor model could partly find explanations from previous research findings of shunned stocks, limited optimization, SRI list inclusion driven demand, individual negatively performing KPIs, or overly optimistic expectations, as discussed in the Section 4. For shunned stocks (Hong and Kacperczyk, 2009), investors pay a financial price by abstaining from investing in socially reprehensible stocks. In the Global 100 portfolio, tobacco and weapons manufacturing companies are excluded, which could explain part of the underperformance. For limited optimization, the relatively harsh initial screen immediately excludes any firm with market capitalization below US$ 2 billion. According to the Markowitz (1959) argument, a constrained optimization is never better than an unconstrained optimization, mathematically. Thus, the exclusionary screen may also partially explain the underperformance. For demand driven by SRI list inclusion, inclusion in the Global 100 list itself could drive demand for stocks on the list, as SRI funds and institutions increase their weights in them. If such an effect is pervasive enough with a significant portion of the investor base acting accordingly, part of the underperformance could be explained by demand-based overly high stock prices. For effects of individual KPIs, it is likely, based on previous research, that some KPIs used in the formation of the Global 100 list might yield positive returns. However, since the net result is negative, there are evidently more KPIs that are value-irrelevant, and drive a portion of the stock underperformance. Lastly, overoptimistic expectations of the benefits ESG performance brings could also partly explain the underperformance. If investors consistently overvalue potential

benefits of ESG factors, but their expectations are met with disappointing tangible results, the stock performance of such firms suffer.

Overall, this paper contributes to the SRI literature by documenting negative risk-adjusted excess stock returns in the global top 100 sustainable firms. The main implication from this finding is that the market overvalues ESG factors, and so investing according to SRI principles faces a financial cost. However, the findings in this paper are limited in establishing any causal link between an isolated factor that causes the underperformance. For future research, the open ends discussed in this paper present ample opportunities for further discoveries.

## References


Aboody, D., and Lev, B. (1998). The value relevance of intangibles: the case of software capitalization. *Journal of Accounting Research, 36,* 161-191.

Aktas, N., de Bodt, E., and Cousin, J.-G. (2011). Do financial markets care about SRI? Evidence from mergers and acquisitions. *Journal of Banking and Finance, 35*(7), 1753-1761.

Bebchuk, L. A., Cohen, A., and Wang, C. C. Y. (2013). Learning and the disappearing association between governance and returns. *Journal of Financial Economics, 108*(2), 323-348.

Benson, K. L., and Humphrey, J. E. (2008). Socially responsible investment funds: Investor reaction to current and past returns. *Journal of Banking and Finance, 32*(9), 1850-1859.

Beiner, S., Drobetz, W., Schmid, M. M., and Zimmermann, H. (2006). An integrated framework of corporate governance and firm valuation. *European Financial Management, 12*(2), 249-283.

Bialkowski, J., and Starks, L. T. (2016). SRI funds: Investor demand, exogenous shocks and ESG profiles. *Working Paper No. 11/2016*. Retrieved from https://ideas.repec.org/p/cbt/econwp/16-11.html

Bollen, N. P. (2007). Mutual fund attributes and investor behavior. *Journal of Financial and Quantitative Analysis, 42*(30), 683-708.

Brammer, S., Brooks, C., and Pavelin, S. (2006). Corporate social performance and stock returns: UK evidence from disaggregate measures. *Financial Management, 35*(3), 97-116.

Carhart, M. M. (1997). On persistence in mutual fund performance. *Journal of Finance, 52*(1), 57-82.

Chan, L., Lakonishok, J., and Sougiannis, T. (2001). The stock market valuation of research and development expenditures. *Journal of Finance, 56,* 2431-2456.

Core, J. E., Guay, R., Rusticus, T. (2006). Does weak governance cause weak stock returns? An examination of firm operating performance and investors' expectations. *Journal of Finance, 61*(2), 655-687.

Cremers, K. J. M., and Nair, V. B. (2005). Governance mechanisms and equity prices. *Journal of Finance, 60*(6), 2859-2894.

Deng, X., Kang, J. K., Low, B. S. (2013). Corporate social responsibility and stakeholder value maximization: Evidence from mergers. *Journal of Financial Economics, 110*(1), 87-109.

Deng, Z., Lev, B., Narin, F. (1999). Science and technology as predictors of stock performance. *Financial Analysts Journal, 55,* 20-32.

Derwall, J., Guenster, N., Bauer, R., and Koedijk, K. (2005). The eco-efficiency premium puzzle. *Financial Analysts Journal, 61*(2), 51-63.

Derwall, J., Koedijk, K., and Ter Horst, J. (2011). A tale of values-driven and profit-seeking social investors. *Journal of Banking and Finance, 35*(8), 2137-2147.

Eccles, R. G., Ioannou, I., and Serafeim, G. (2014). The impact of corporate sustainability on organizational processes and performance. *Management Science, 60*(11), 2835-2857.

Edmans, A. (2011). Does the stock market value intangibles? Employee satisfaction and equity prices. *Journal of Financial Economics, 101*, 621-640.

Fama, E. F., and French, K. R. (1997). Industry costs of equity. *Journal of Financial Economics, 43*(2), 153-193.

Fama, E. F., and French, K. R. (2008). Dissecting anomalies. *Journal of Finance, 63*(4), 1653-1678.

Fama, E. F., and French, K. R. (2016). Dissecting anomalies with a five-factor model. *Review of Financial Studies, 29*(1), 69-103.

Flammer, C. (2015). Does corporate social responsibility lead to superior financial performance? A regression discontinuity approach. *Management Science, 61*(11), 2549-2568.

Galema, R., Plantinga, A., and Scholtens, B. (2008). The stocks at stake: Return and risk in socially responsible investment. *Journal of Banking and Finance, 32*(12), 2646-2654.

Gompers, P. A., Ishii, J. L., and Metrick, A. (2003). Corporate governance and equity prices. *Quarterly Journal of Economics, 118*(1), 107-155.

Heinkel, R., Kraus, A., and Zechner, J. (2001). The effect of green investment on corporate behavior. *Journal of Financial and Quantitative Analysis, 36*(4), 431-449.

Hong, H., and Kacperczyk, M. (2009). The price of sin: The effects of social norms on markets. *Journal of Financial Economics, 93*(1), 15-36.

Kempf, A., and Osthoff, P. (2007). The effect of socially responsible investing on portfolio performance. *European Financial Management, 13*(5), 908-922.

Lev, B., and Sougiannis, T. (1996). The capitalization, amortization, and value-relevance of R&D. *Journal of Accounting and Economics, 21*, 107-138.

Margolis, J. D., Elfenbein, H. A., and Walsh, J. P. (2008). Do well by doing good? Don't count on it. Social Responsibility. Special Issue on HBS Centennial. *Harvard Business Review 86*(1).

Markowitz, H. (1959). Portfolio Selection: Efficient Diversification of Investments. Wiley, New York.

Merton, R. (1987). A simple model of capital market equilibrium with incomplete information. *Journal of Finance, 42*, 483-510.

Newey, W. K., and West, K. D. (1987). Hypothesis testing with efficient method of moments estimation. *International Economic Review*, 777-787.

Nofsinger, J. and Varma, A. (2014). Socially responsible funds and market crises. *Journal of Banking and Finance, 48*, 180-193.

Orlitzky, M., Schmidt, F. L., and Rynes, S. L. (2003). Corporate social and financial performance: A meta-analysis. *Organization Studies, 24*(3), 403-441.

Renneboog, L., Ter Horst, J., and Zhang, C. (2008). The price of ethics and stakeholder governance: The performance of socially responsible mutual funds. *Journal of Corporate Finance, 14*(3), 302-322.

Statman, M., and Glushkov, D. (2009). The wages of social responsibility. *Financial Analysts Journal, 65*(4), 33-46.